Supporting Information

A network of Spies: synthesis of bioactive protein hydrogels by genetically-encoded SpyTag-SpyCatcher chemistry

Fei Sun*[†], Wen-Bin Zhang*[†], Alborz Mahdavi*, Frances H. Arnold*[‡] and David A. Tirrell*[‡]

*Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States

[†]F.S. and W.-B.Z. contributed equally to this work.

[‡]To whom correspondence should be addressed:

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States

Phone: 626-3954162; Email: frances@cheme.caltech.edu

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States

Phone: 626-3953140; Email: tirrell@caltech.edu

Supporting Figures

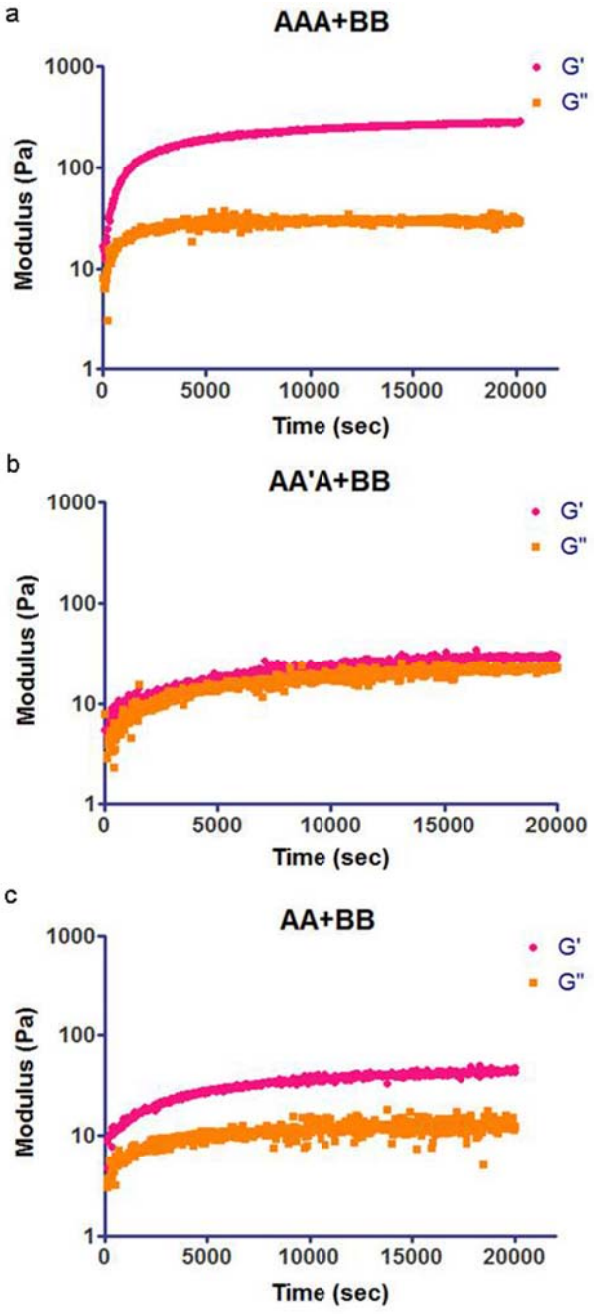

Fig. S1. Evolution of storage (G') and loss (G") moduli as a function of time during *in situ* gelation at 25°C. Shear frequency and strain amplitude were held constant at 1 rad/sec and 5%, respectively.

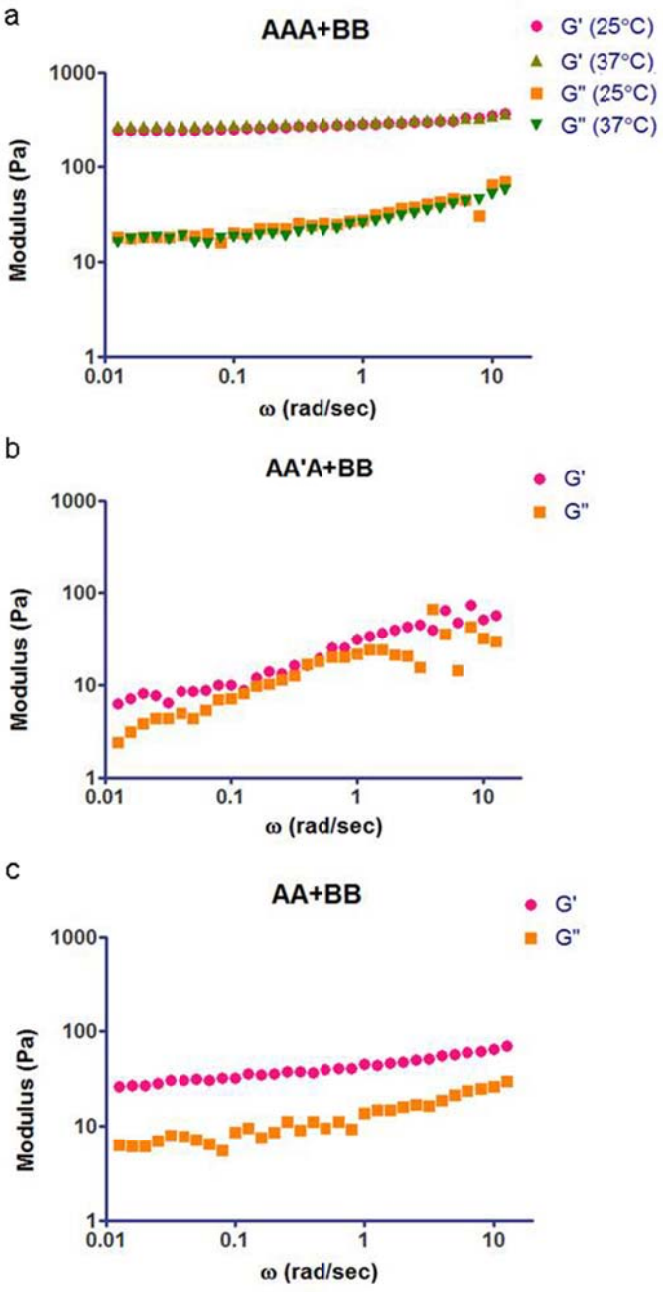

Fig. S2. Frequency sweep tests on AAA+BB (*a*), AA'A+BB (*b*) and AA+BB (*c*). Measurements on AAA+BB (10 wt%) were performed at both 25 and 37°C. Measurements on AA'A+BB (10 wt%) and AA+BB (10 wt%) were done at 25°C. Shear frequency decreased from 100 to 0.01 rad/sec and strain was fixed at 5%. Both storage and loss moduli, G' and G", were recorded. Data are shown only for the frequency range 12.6 to 0.01 rad/sec; measurements of G" at higher frequencies were characterized by high noise levels.

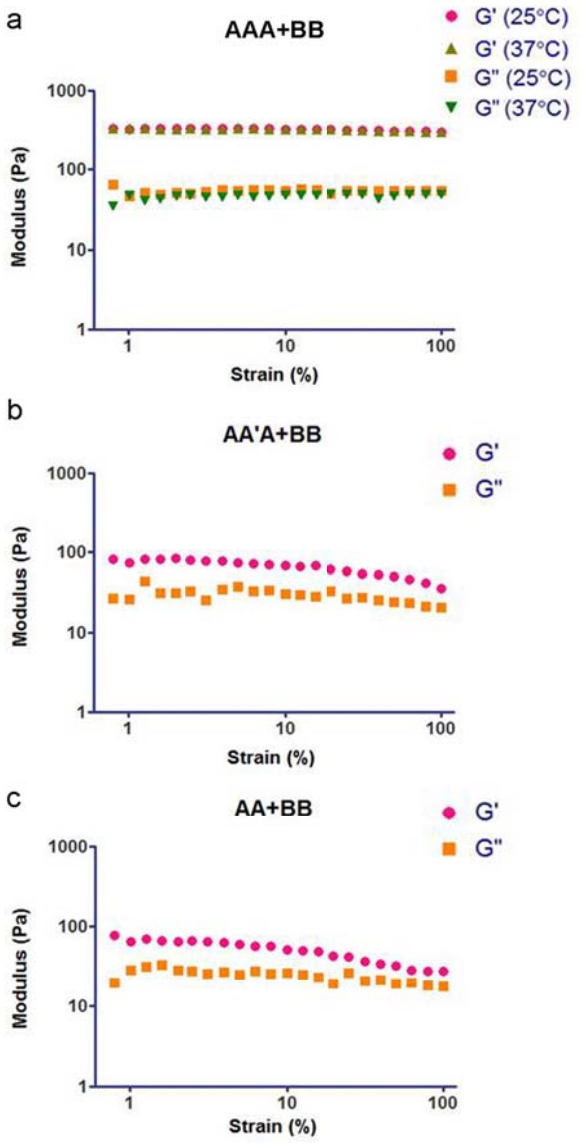

Fig. S3. Strain sweep tests on AAA+BB (*a*), AA'A+BB (*b*) and AA+BB (*c*). Measurements on AAA+BB (10 wt%) were performed at both 25 and 37°C. Measurements on AA'A+BB (10 wt%) and AA+BB (10 wt%) were done at 25°C. Shear strain increased from 0.1% to 100%; frequency was fixed at 10 rad/sec. Both storage and loss moduli, G' and G", were recorded. Data acquired at strain amplitudes below 0.8% were characterized by high noise levels.

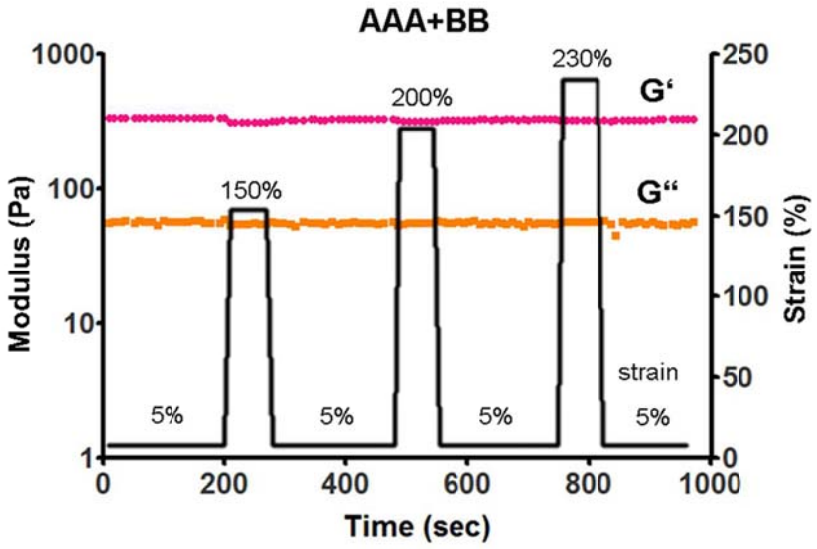

Fig. S4. Large-amplitude oscillatory shear tests of the Spy network at 10 rad/sec.

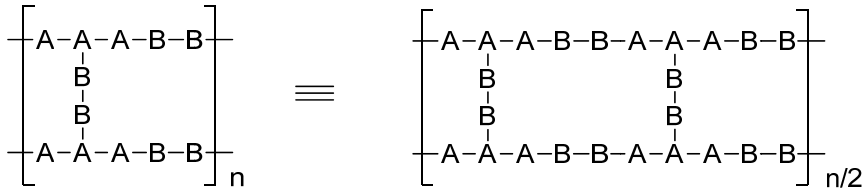

Fig. S5. Schematic of a 100% crosslinked Spy network comprised of AAA and BB at a molar ratio of 2:3. The molecular weight between crosslinks, M_c, equals $(2M_{AAA}+3M_{BB})/2$.

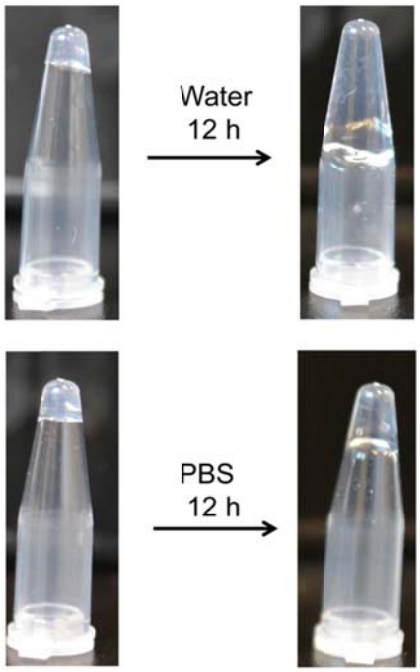

Fig. S6. Swelling of the Spy network (AAA+BB) in water (*top*) and PBS buffer (*bottom*).

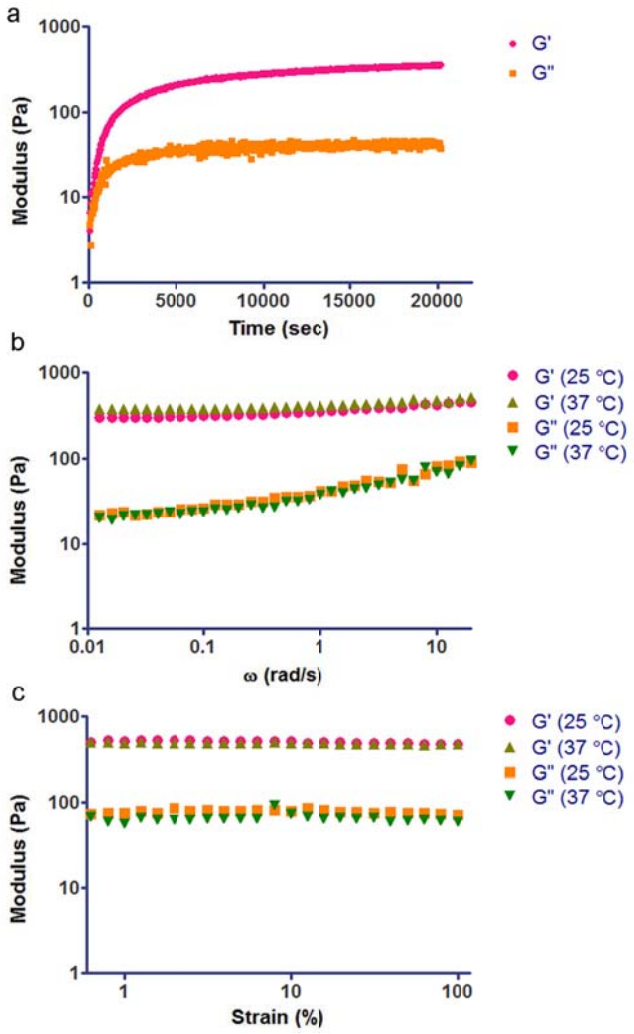

Fig. S7. Rheology of mCherry-Spy-network. (*a*) Time sweep tests of *in situ* gelation of mCherry-Spy-network at 25°C. The shear frequency was held constant at 1 rad/sec; the strain amplitude at 5%. (*b*) Frequency sweep tests on 6 hour-cured mCherry-Spy-network at 25°C and 37°C. The shear strain amplitude was set at 5%. The shear frequency varied from 100 to 0.01 rad/sec. Measurements of G" at frequencies above 20 rad/sec were characterized by high noise levels. (*c*) Strain sweep tests on 6 hour-cured mCherry-Spy-network at 25°C and 37°C. The shear frequency was set at 10 rad/sec. The shear strain increased from 0.1 to 100%. Measurements of G" at strain amplitudes below 0.5% were characterized by high noise levels.

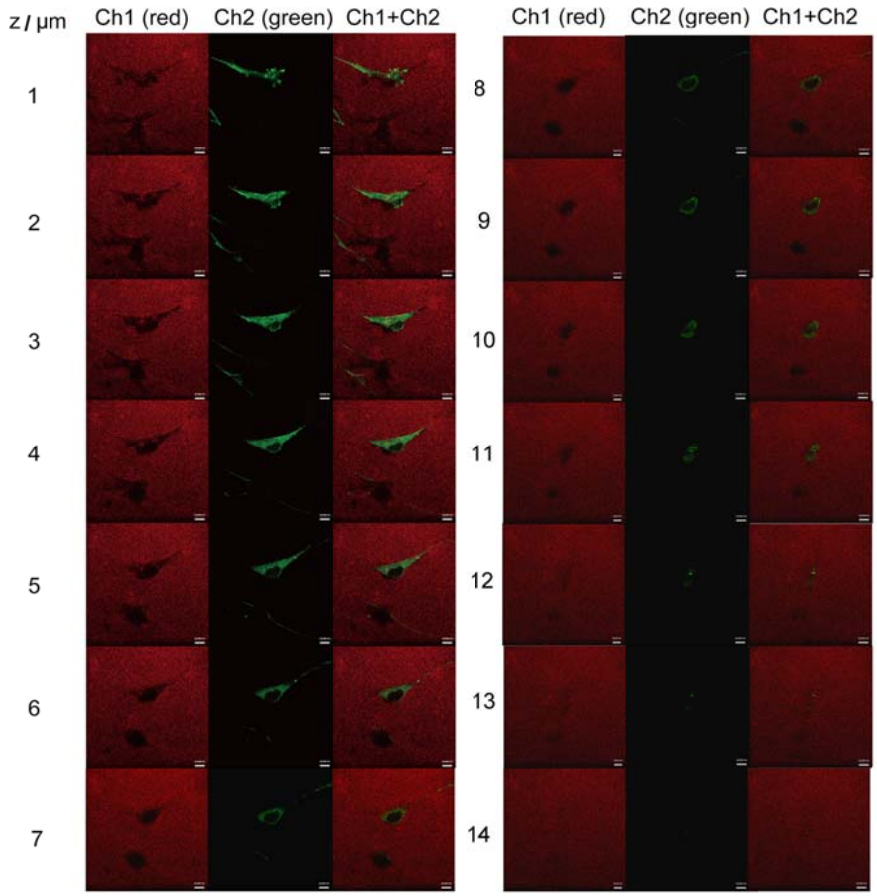

Fig. S8. Confocal fluorescence z-slice micrographs of mouse NIH 3T3/GFP fibroblasts encapsulated in the mCherry-Spy-network. Channel 1 (Ch1) images reveal the mCherry-Spy-network (red), in which dark areas correspond to regions occupied by 3T3 cells. Channel 2 (Ch2) images show two spreading 3T3/GFP cells (green). Ch1+Ch2 images are overlays of the matrix and cells (Scale bars, 10 µm).

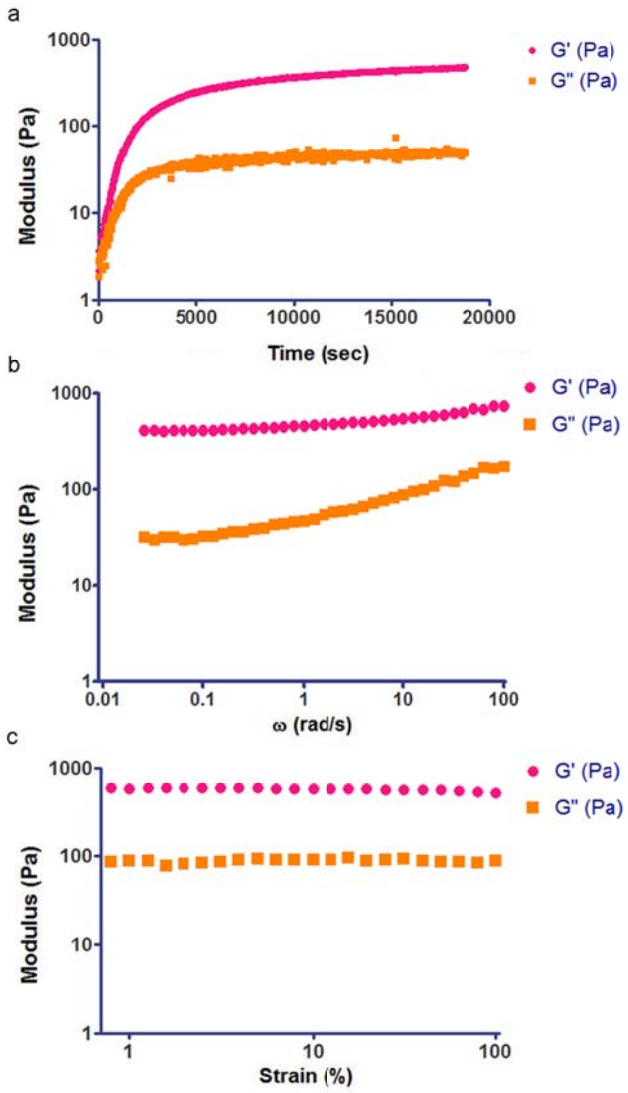

Fig. S9. Rheology of LIF-Spy-network. (*a*) Time sweep tests of *in situ* gelation of LIF-Spy-network at 25°C. AAA (10 wt%) was dissolved in an A-LIF-A solution in distilled water. BB (10 wt%) was dissolved in mESC medium. Gelation was initiated by mixing the two precursors in a molar ratio of 2:3 (AAA:BB). Shear frequency was held constant at 1 rad/sec; strain amplitude at 5%. (*b*) Frequency sweep tests of 6 hour-cured LIF-Spy-network at 25°C. The shear strain amplitude was set at 5%. The shear frequency varied from 100 to 0.025 rad/sec. (*c*) Strain sweep tests of 6 hour-cured LIF-Spy-network at 25°C. The shear frequency was set at 10 rad/sec. The shear strain amplitude varied from 0.1 to 100%. Measurements of G" at strain amplitudes below 0.8% were characterized by high noise levels.

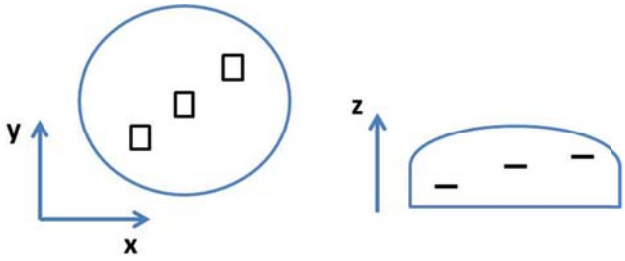

Fig. S10. Illustration of bright-field images acquired from different regions of each Spy-network gel.

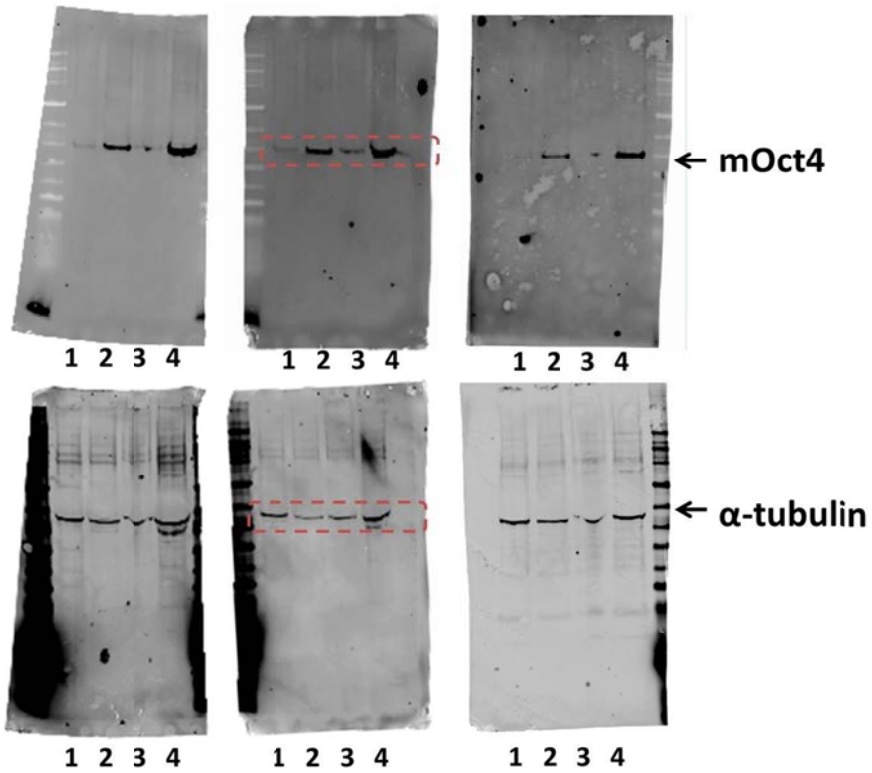

Fig. S11. Full-scan images from immunoblotting experiments.

a. AAA

MKGSSHHHHH HVDAHIVMVD AYKPTKLDGH GVGVPGVGVP GVGVPGEGVP GVGVPGVGVP
GVGVPGVGVP GEGVPGVGVP GVGVPGVGVP GVGVPGEGVP GVGVPGVGEL AHIVMVDAYK
PTKTSVPGVG VPGVGVPGEG VPGVGVPGVG VPGVGVPGVG VPGEGVPGVG VPGVGVPGVG
VPGVGVPGEG VPGVGVPGVG VPGGLLDAHI VMVDAYKPTK LEWKK

b. AA'A

MKGSSHHHHH HVDAHIVMVD AYKPTKLDGH GVGVPGVGVP GVGVPGEGVP GVGVPGVGVP
GVGVPGVGVP GEGVPGVGVP GVGVPGVGVP GVGVPGEGVP GVGVPGVGEL AHIVMVAAYK
PTKTSVPGVG VPGVGVPGEG VPGVGVPGVG VPGVGVPGVG VPGEGVPGVG VPGVGVPGVG
VPGVGVPGEG VPGVGVPGVG VPGGLLDAHI VMVDAYKPTK LEWKK

Fig. S12. Amino acid sequences of (*a*) AAA and (*b*) AA'A. Reactive Asp residues are shown in the boxes. The segments in red are SpyTags. In AA'A, the reactive Asp residue is mutated to Ala (grey). The sequences of AA and BB have been published (1).

A-mCherry-A

MKGSSHHHHH HVDAHIVMVD AYKPTKLDGH GVGVPGVGVP GVGVPGEGVP GVGVPGVGVP
GVGVPGVGVP GEGVPGVGVP GVGVPGVGVP GVGVPGEGVP GVGVPGVGEL VSKGEEDNMA
IIKEFMRFKV HMEGSVNGHE FEIEGEGEGR PYEGTQTAKL KVTKGGPLPF AWDILSPQFM
YGSKAYVKHP ADIPDYLKLS FPEGFKWERV MNFEDGGVVT VTQDSSLQDG EFIYKVKLRG
TNFPSDGPVM QKKTMGWEAS SERMYPEDGA LKGEIKQRLK LKDGGHYDAE VKTTYKAKKP
VQLPGAYNVN IKLDITSHNE DYTIVEQYER AEGRHSTGGM DELYKTSVPG VGVPGVGVPG
EGVPGVGVPG VGVPGVGVPG VGVPGEGVPG VGVPGVGVPG VGVPGVGVPG EGVPGVGVPG
VGVPGGLLDA HIVMVDAYKP TKLEWKK

Fig. S13. Amino acid sequence of A-mCherry-A. Reactive Asp residues are shown in the boxes. The segments in red and in yellow are SpyTags and mCherry, respectively.

A-LIF-A

```
MKGSSHHHHH HVDAHIVMVD AYKPTKLDGH GVGVPGVGVP GVGVPGEGVP GVGVPGVGVP   60
GVGVPGVGVP GEGVPGVGVP GVGVPGVGVP GVGVPGEGVP GVGVPGVGEL SPLPITPVNA  120
TCAIRHPCHG NLMNQIKNQL AQLNGSANAL FISYYTAQGE PFPNNLDKLC GPNVTDFPPF  180
HANGTEKAKL VELYRMVAYL SASLTNITRD QKVLNPSAVS LHSKLNATID VMRGLLSNVL  240
CRLCNKYRVG HVDVPPVPDH SDKEVFRKKK LGCQLLGTYK QVISVVVQAF TSVPGVGVPG  300
VGVPGEGVPG VGVPGVGVPG VGVPGVGVPG EGVPGVGVPG VGVPGVGVPG VGVPGEGVPG  360
VGVPGVGVPG GLLDAHIVMV DAYKPTKLEW KK                               392
```

Fig. S14. Amino acid sequence of A-LIF-A. Reactive aspartic acid residues are shown in the boxes. The segments colored in yellow and in red are LIF and SpyTag, respectively.

Supporting Discussion

Calculation of crosslink densities from storage modulus G' and swelling ratio *q*:

a. Molecular weight between crosslinks of the perfect Spy network (100% crosslinking):

   $M_c = (3M_{BB}+2M_{AAA})/2 = 91.5$ kDa

b. Molecular weight between crosslinks deduced from the storage modulus:

   $M_c' = \rho RT/G = 826$ kDa

where ρ is the polymer density (0.100 g/cm^3 for a 10 wt% network) and G is the storage modulus (0.3 kPa).

The estimated crosslinking efficiency (M_c/M_c') = 11%.

c. Molecular weight between crosslinks deduced from the swelling ratio in water:

   From Flory network theory (2):

$$M_c = \frac{V_1\rho_2\left(V_{2m}^{1/3} - \frac{V_{2m}}{2}\right)}{-[\ln(1 - V_{2m}) + V_{2m} + \chi_1 V_{2m}^2]}$$

where $V_{2m} = q^{-1}$ (q = swelling ratio); ρ_2 is the polymer density after swelling in water (3.2×10^{-3} g/cm^3); V_1 is the molar volume of solvent (18 cm^3/mol); and χ_1 is the polymer-solvent interaction parameter. For hydrophilic polymers such as polyacrylamide, poly(N,N-dimethylacrylamide), poly(N-isopropylacrylamide), and poly(vinyl alcohol) in water at 25 °C, $\chi_1 = \sim 0.5$ (3). Although χ_1 is not known for the proteins that form the Spy network, we used a value of 0.5 to estimate M_c from the swelling data. In distilled water, the swelling ratio, q, is 310, and a value of 0.5 for χ_1 gives $M_c = \sim$750 kDa, in reasonable agreement with the value of M_c calculated from the storage modulus.

The crosslink densities derived from the storage modulus and the swelling ratio are approximately ten-fold lower than that of the perfect Spy network, indicating incomplete crosslinking.

Table S1. Bacterial strains，plasmids, and primers used in this study.

Strains	**Relevant Characteristics**	**Source**
E. coli		
DH5α		Stratagene
BL21 star(DE3)		Invitrogen
Plasmids	**Relevant Characteristics**	**Source**
pQE-80L	T5 promoter-operator, N-terminal His tag, Amp^r	Qiagen
pQE-*ELP*	pQE-80L plasmid containing the gene encoding elastin	Starting vector for all recombinant proteins in this study
pQE-*EA*	Plasmid for expression of Elastin-SpyTag	This study
pQE-*EB*	Plasmid for expression of Elastin-SpyCatcher	This study
pQE-*AA*	Plasmid for expression of SpyTag-Elastin-RGD-Elastin-SpyTag	This study
pQE-*BB*	Plasmid for expression of SpyCatcher-Elastin-RGD-Elastin-SpyCatcher	This study
pQE-*AAA*	Plasmid for expression of SpyTag-Elastin-SpyTag-Elastin-SpyTag	This study
pQE-*AA'A*	pQE-AAA-D117A mutant, where the internal SpyTag is inactivated by Asp117 to Ala mutation	This study
pQE-*A-mCherry-A*	Plasmid for expression of SpyTag-Elastin-mCherry-Elastin-SpyTag	

Primers	**Sequence**
SpyCatcher-*Sal*I-F	GTGTACA GTCGAC A TC CCA ACG AC C GAA AAC CTG TATTTTC
spyCatcher-*Xho*I-R	GTAGGCACTCGAG CTGACCCCAAATACCTTGCGGACCGTC
SpyTag-*Sal*I-*Xho*I-F	GACA GTCGACGCCCATATTGTCATGGTTGATGCATACAA GCCGACGAAGCTCGAG GTAGG
SpyTag-*Sal*I-*Xho*I-R	CCTACCTCGAGCTTCGTCGGCTTGTATGCATCAACC ATGACAATATGGGCGTCGACTGTC
SpyTag-*Sac*I-*Spe*I-F	GACA GAGCTCGCCCATATTGTCATGGTTGATGCATACAA GCCGACGAAGACTAGT GTAGG
SpyTag-*Sac*I-*Spe*I-R	CCTACACTAGTCTTCGTCGGCTTGTATGCATCAACC ATGACAATATGGGCGAGCTCTGTC
SpyTagD20A-*Sac*I-*Spe*I-F	GACA GAGCTC GCCCATATTGTCATGGTTGCGGCATACAAGCCGAC GAAGACTAGT GTAGG
SpyTagD20A-*Sac*I-*Spe*I-R	CCTACACTAGTCTTCGTCGGCTTGTATGCCGCAACC ATGACAATATGGGCGAGCTCTGTC
mCherry-*Sac*I-F	ATAA GAGCTC ATGGTGAGCAAGGGCGAGG
mCherry-*Spe*I-R	CAA TAC TAG TCT TGT ACA GCT CGT CCA TGC CGT C
LIF-*Sac*I-F	GTAC GTCGAC GAGCTC TCT CCG CTG CCG ATC ACC CCG GTA AAC GCA ACC TGT GCC ATC C
LIF-*Spe*I-R	GTAGCTCGAGACTAGTAAACGCCTGAACAACCACA GAGATAACCTGTTTGTAAGTAC

References

1. Zhang WB, Sun F, Tirrell DA, & Arnold FH (2013) Controlling macromolecular topology with genetically encoded SpyTag-SpyCatcher chemistry. *J Am Chem Soc* 135(37):13988-13997.

2. Sperling LH (2006) *Introduction to Physical Polymer Science*, 4th edition, pp. 472-473. Wiley & Sons. ISBN: 978-0-471-70606-9.

3. Orwell RA and Arnold PA (2006) Chapter 14. Polymer-Solvent Interaction Parameter χ.In Mark JE (ed.) *Physical Properties of Polymers Handbook*, 2nd edition, pp. 233-256. Springer. ISBN: 978-0-387-31235-4.